

REMARKS

Reexamination and reconsideration of this application as amended is requested.
By this amendment, Claims 1, 4, 7-9, 11, 13, 16, and 19-20, have been amended, and Claims 3 and 15 have been canceled. After this amendment, Claims 1-2, 4-14, and 16-20, remain pending in this application.

Allowable Subject Matter

The Examiner objected to Claims 8 and 20, but indicated these claims would be allowable if rewritten in independent form including all limitations of the base claim and any intervening claims.

Applicants wish to acknowledge and thank the Examiner for recognizing the allowable subject matter in these claims, and accordingly have amended Claims 8 and 20 to comply with Examiner's request. These claims now recite in independent form including all limitations of the base claim and any intervening claims.

Therefore, Applicants submit that Claims 8 and 20 are in allowable form, and kindly request that the Examiner allow these claims to issue.

Claims Rejection under 35 U.S.C. § 112, second paragraph

(4-7) The Examiner rejected Claims 4-6, and 16-18, under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Specifically, the Examiner refers to the exemplary process discussed in the specification at page 13, line 10, through page 14, line 4, as being inconsistent because the Examiner indicated that in order for a node to be determined a pagelet all the children need to undergo this analysis, and these children nodes are not inserted into the queue until after it has been determined that the parent

node is not a pagelet.

Applicants wish to clarify a possible confusion with respect to the discussion in the specification and from the Examiner's comments. The test at step 612, and as specifically discussed on page 13, lines 16-17, and on page 14, lines 1-4, addresses when in the affirmative the node is determined to be a pagelet. If the node *v* is declared a pagelet, it is output at step 616. However, if it can not be determined at this pass that the node is a pagelet (i.e, see the "Otherwise" on page 14, line 2) then all of the children nodes of the parent node are inserted into the queue at step 614. The process is repeated until the queue is empty. Note for example that the queue in the second pass may include the first parent node *v* (that could not be immediately determined a pagelet) and its children nodes and then processing should continue with each of the children nodes in the queue to determine whether any are pagelets. This process keeps inserting children nodes of a parent node that can not be immediately determined to be a pagelet and then processing of these children nodes in the queue continues until all pagelet nodes are outputted and the queue is empty. This is merely just one example of traversing a hypertext parse tree 422 for identifying pagelet nodes. Those of ordinary skill in the art will appreciate many other ways to traverse such a tree for finding pagelet nodes in view of the discussion in the specification since the general process of traversing trees is well known to practitioners of ordinary skill in the relevant art.

In view of the remarks above, Applicants believe that the discussion in the specification for providing an example of traversing a hypertext parse tree should be better understood, and Applicants kindly request that the Examiner withdraw the rejection of Claims 4-6 and 16-18, under 35 U.S.C. § 112, second paragraph. Applicants also welcome the Examiner to call our undersigned representative if further clarification and/or discussion of this rejection would help in any way expedite prosecution.

Claim Rejections - under 35 USC § 102

(8-16) The Examiner rejected Claims 1-3, 7, 9, 11, 13-15 and 19 under 35 U.S.C. 102(b) as being anticipated by Bröder et al. ("Syntactic Clustering of the Web").

Applicants have amended independent Claims 1, 9, 11, and 13, to more clearly and distinctly recite the present invention. Additionally, note that dependent Claims 3 and 15 were canceled and their limitations were added to their respective independent Claims 1 and 13. The amended independent claims, and all dependent claims depending therefrom respectively, now more clearly and distinctly recite that a template comprises a collection of pagelets T satisfying the following two requirements: (1) all the pagelets in T are identical or almost identical; and (2) every two pages owning pagelets in T are reachable one from the other via other pages also owning pagelets in T. This definition of a template in the independent claims should clearly distinguish the presently claimed "eliminating the template pagelets" from the teachings in Broder, section 5.1, regarding document resemblance comparisons while ignoring common shingles that are defined by Broder as nearly all mechanically generated and including shared header or footer information on a large number of automatically generated pages, i.e., forms or views on databases. While the Office Action previously had mischaracterized the claimed templates as being analogous to such forms or views on databases, i.e., the common shingles, it should be clear now that the claimed templates (a collection of pagelets meeting the two requirements) and further the claimed processing of templates are different than the teachings in Broder, i.e., the common shingles.

Applicants refer the Examiner to FIG. 11 and the specification on page 14, lines 6-16, regarding the definition and example of a template with respect to the presently claimed invention. Note that as a first condition all of the template pagelets are identical or almost identical. See FIG. 11 for examples. As a second condition, the two pages containing template pagelets must be reachable one from the other possibly via other

pages also owning template pagelets in the collection. This second condition regarding the presently claimed template pagelet is not taught or suggested by Broder's discussion of common shingles including forms or views into databases. The processing of template pagelets as presently claimed, and as illustrated by example in FIG. 11, should be understood as being different processing from Broder's ignoring of common shingles including forms and views into databases as part of comparing resemblance of two pages.

Support for this amendment may be found in the specification as originally filed, see for example on page 14, lines 6-10. No new matter was added.

Therefore, in view of the amendment and remarks above, Applicants believe that since Broder does not teach, anticipate, or suggest, the presently claimed invention, the rejection of Claims 1-3, 7, 9, 11, 13-15 and 19 under 35 U.S.C. 102(b) has been overcome. The Examiner should withdraw the rejection of these claims.

Claim Rejections - 35 USC § 103

(17-21)The Examiner rejected Claims 4 and 16 under 35 U.S.C. 103(a) as being unpatentable over Broder et al. ("Syntactic Clustering of the Web") as applied to claims 1-3, 7, 9, 11, 13-15 and 19 and further in view of Chakrabarti et al. ("Enhanced Topic Distillation using Text, Markup Tags and Hyperlinks").

Applicants have amended dependent Claims 4 and 16 to depend from independent Claims 1 and 13, respectively, and the dependent claims recite all of the limitations of the independent claims. In view of the discussion above, Applicants believe that it should be clear that the claimed processing of template pagelets, as recited for dependent Claims 4 and 16, is not taught or suggested by Broder. While Chakrabarti was included in the instant rejection to arguably obviate the detailed decomposing step not taught or suggested by Broder, Applicants wish to point out that Chakrabarti's processing is different than the presently claimed decomposing step and further that there is no teaching or suggestion for

the presently claimed eliminating the template pagelets. It is believed that in view of the amendment and remarks above, neither cited reference nor any combination thereof, teaches or suggests the presently claimed invention and dependent claims 4 and 16 recite in allowable form. The Examiner should withdraw the rejection of these claim.

(22-23) The Examiner rejected Claims 10 and 12 under 35 U.S.C. 103(a) as being unpatentable over Broder et al. ("Syntactic Clustering of the Web") as applied to claims 1-3, 7, 9, 11, 13-15 and 19 and further in view of Rodeheffer et al. (U.S. Patent No. 6,614,764). Specifically, the Examiner cited the combination of Broder with the Rodeheffer reference to arguably obviate the claimed invention specifically with respect to the Breadth First Search (BFS) algorithm.

Applicants have amended independent Claims 9 and 11 to more clearly and distinctly recite the present invention. Amended Claims 9 and 11 now more clearly and distinctly recite the eliminating the template pagelets wherein a template comprises a collection of pagelets T satisfying the following two requirements: (1) all the pagelets in T are identical or almost identical; and (2) every two pages owning pagelets in T are reachable one from the other via other pages also owning pagelets in T.

As has already been discussed above, a template as used in the independent Claims 9 and 11, and as defined in the specification of the present patent application, should clearly distinguish the presently claimed "eliminating the template pagelets" from the teachings in Broder, section 5.1. Broder's teachings in section 5.1 are directed at document resemblance comparisons while ignoring common shingles that are defined by Broder as nearly all mechanically generated and including shared header or footer information on a large number of automatically generated pages, i.e., forms or views on databases. This is different than the presently claimed templates and processing of templates, as recited for amended Claims 9 and 11.

The Examiner characterized Rodeheffer as teaching the Breadth First Search

(BFS) technique. However, Rodeheffer does not teach or suggest the presently claimed templates and template processing, as recited for independent Claims 9 and 11. Accordingly, it should be clear that Broder, Rodeheffer, or any combination of the two cited references, does not teach or suggest the presently claimed invention, as recited for independent Claims 9 and 11 as discussed above.

Dependent Claims 10 and 12 depend from amended independent Claims 9 and 11, respectively. As discussed above, amended independent Claims 9 and 11 distinguish over the cited prior art. Since dependent claims include all of the limitations of the independent claims from which they depend, Applicants further assert that dependent Claims 10 and 12 also distinguish over the cited prior art as well. Therefore, Applicants respectfully assert that neither Broder, Rodeheffer, nor any combination of the two cited references, teaches or suggests the presently claimed invention as recited for dependent Claims 10 and 12, and the Examiner's rejection of Claims 10 and 12 under 35 U.S.C. §103(a) should be withdrawn.

Allowable Subject Matter

(24) As has already been discussed above, Applicants have amended Claims 8 and 20 to recite in independent form including all limitations of the base claim and any intervening claims to comply with Examiner's request.

Therefore, Applicants believe that Claims 8 and 20 are in allowable form, and kindly request that the Examiner allow these claims to issue.

Conclusion

The foregoing is submitted as full and complete response to the Official Action mailed May 24, 2004, and it is submitted that Claims 1-2, 4-14, and 16-20 are in condition for allowance. Reconsideration of the rejection is requested. Allowance of Claims 1-2, 4-14, and 16-20 is earnestly solicited.

No amendment made was related to the statutory requirements of patentability unless expressly stated herein. No amendment made was for the purpose of narrowing the scope of any claim, unless Applicants have argued herein that such amendment was made to distinguish over a particular reference or combination of references.

Applicants acknowledge the continuing duty of candor and good faith to disclosure of information known to be material to the examination of this application. In accordance with 37 CFR §§ 1.56, all such information is dutifully made of record. The foreseeable equivalents of any territory surrendered by amendment are limited to the territory taught by the information of record. No other territory afforded by the doctrine of equivalents is knowingly surrendered and everything else is unforeseeable at the time of this amendment by the Applicants and the attorneys.

If the Examiner believes that there are any informalities that can be corrected by Examiner's amendment, or that in any way it would help expedite the prosecution of the patent application, a telephone call to the undersigned at (561) 989-9811 is respectfully solicited.

Claims Amendment Fee

The present application, after entry of this amendment, comprises eighteen claims (18) claims, including six (6) independent claims. Applicants have previously paid for twenty (20) claims including four (4) independent claims. Applicants, therefore, believe that an additional fee (\$176 = 2 X \$88) for two (2) additional independent claims is currently due. The Commissioner is hereby authorized to charge the fee of \$176 for claims amendment, or if this fee amount is insufficient or incorrect then the Commissioner is authorized to charge the appropriate fee amount to prevent this application from becoming abandoned, or credit any overpayment, to Deposit Account 09-0441.